Through Thick and Thin: MRI Super-Resolution Using a Generative Adversarial Network

Andrew Narcomey
aon1@stanford.edu

Angela Gu
angelajg@stanford.edu

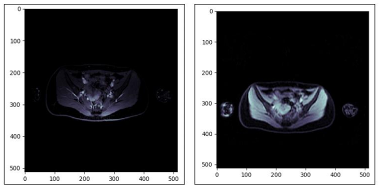
Objective

- Because MRI (Magnetic Resonance Imaging) does not emit radiation, it is the preferred scanner for soft tissue differentiation, frequent imaging, or vulnerable patients.
- However, scans require patients to lie still inside the scanner for half an hour or more, which may be difficult.
- Therefore, there may be a trade-off between the quality of the scan and the patient's comfort.

Can we use deep learning to reconstruct high-resolution images from rapidly acquired low-resolution data?

Data

- This dataset was provided by Dr. Anuj Pareek (Postdoc in Radiology).
- It contains paired high- and low-resolution scans from 16 patients. Each scan is composed of several hundred "slices" of varying thickness.
- Images were *dicom* files: an array of pixels (512 x 512) with 1 channel, as well as metadata about the scan and the patient.



LR MRIs scan less thoroughly and, thus, have thicker and fewer slices. We used the metadata to pair each HR scan (left) with its best-fitting LR scan (right).

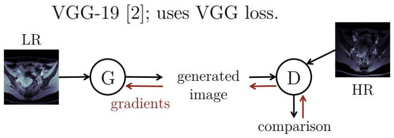
Image super-resolution techniques typically focus on minimizing MSE (mean-squared error) between the pixels of the two images:

$$MSE(I, \hat{I}) = \frac{1}{n} \sum_{i,j} (I_{i,j} - \hat{I}_{i,j})^2$$

However, MSE can be very sensitive to small shifts in the image, and our paired images showed noticeable differences in the pixel values and placement.

Models

- SRGAN [1] defines the *VGG loss* as the Euclidean distance between the feature representations of a reconstructed image and its reference high-resolution image.
- Using SRGAN as reference, we trained:
  - Generator *G*: residual network (SRResNet); uses MSE loss.
  - Discriminator *D*: similar architecture to VGG-19 [2]; uses VGG loss.



Modifications include: creating paired images (LR and HR) instead of down-sampling HR images; adapting all parameters in *G*, *D*, and *VGG-19* for 1-channel images

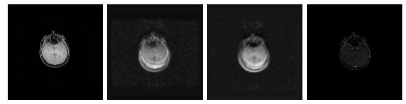
MRIs have an additional feature in that they are sequential; therefore, a slice's neighbors may contain useful information. Therefore, we run a similar model, but with our "ground truth" input as a 3-channeled image.

[1] ran SRResNet over 100 iterations and SRGAN over 2000 iterations, which took several weeks. We did not have the computational resources to do this, so all results are in a preliminary stage.

Results & Discussion

Model 1 (without neighbors)

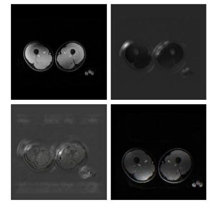
Training: 11 patients; Test: 5 patients
SSResNet: 1 epoch; SRGAN: 3 epochs

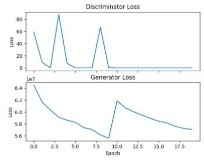


LR SRResNet SRGAN HR

Model 2 (with neighbors)

Training: 2 patients; Test: 1 patient (restricted by storage)
SSResNet: 5 epochs; SRGAN: 20 epochs





LR SRResNet
SRGAN HR

Discussion

For Model 1, SRGAN seems to perform better than SRResNet, while the opposite is true for Model 2—perhaps the small shifts in the slices just add unnecessary noise, instead of useful information. We can also see that (intuitively) more training leads to better results, as seen in Model 2's SRResNet.

We originally intended our evaluation metric to be MOS (Mean Opinion Score) by radiologists. However, given the quality of the generated image, this is not yet possible.

Conclusion

- We were *heavily* restricted by computational resources, particularly access to GPUs, storage, and training time.
- To determine how effective our model is for image super-resolution would require much longer training.
- Nonetheless, our preliminary results for SRResNet and SRGAN show an opportunity for improvement, and we are excited to see results after more extensive testing.

Future Work

- Increased computational power, storage, and time training. Due to patient privacy issues, we were heavily restricted in our access to GPUs.
- Other ways to incorporate the sequential nature of MRIs: recurrent neural networks, 3D CNNs, research into video interpolation techniques.
- Start the algorithm with image detection in order to locate the relevant scan area. This would allows us to reduce the dimensions of the image and thus reduce the computations needed.

References

Thank you to Dr. Anuj Pareek for his contributions!

[1] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A.P., Tejani, A., Totz, J., Wang, Z. and Shi, W., 2017, July. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In CVPR (Vol. 2, No. 3, p. 4).
[2] Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.1556.